UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK

In Re: OpenAI, Inc. Copyright Infringement Litigation	1:25-md-03143-SHS
	This Document relates to:
AUTHORS GUILD, et al., Plaintiffs, v. OPENAI, INC., et al., Defendants.	Case No. 1:23-cv-08292-SHS-OTW
JONATHAN ALTER, et al. Plaintiff, v. OPENAI, INC., et al., Defendants.	Case No. 1:23-cv-10211-SHS-OTW

DEFENDANT MICROSOFT CORPORATION'S OPPOSITION TO NY CLASS PLAINTIFFS' MOTION FOR LEAVE TO AMEND

TABLE OF CONTENTS

TABLE OF AUTHORITIES

Page(s)

Cases

Rules

Other Authorities

INTRODUCTION

A few days after the Judicial Panel on Multidistrict Litigation created the conditions to bring these important cases to efficient resolution, the NY Author Plaintiffs seek to upend the apple cart with their proposed First Amended Consolidated Complaint ("Proposed FACC"). Just as document discovery has begun to wind down and depositions are poised to commence—all of which required frequent and significant attention by this Court—the NY Author Plaintiffs want to start over with new parties, additional claims, and brand-new models and products about which zero discovery has been taken.

It is not surprising that the technology at issue in this case is developing rapidly. While the standard for amending is permissive, it is not designed to encourage a hamster wheel of discovery where more and more and more products are swept in without ever getting closer to a resolution. The Proposed FACC disregards the rationale and demands of centralization and would unduly prolong this matter.

In particular, the Proposed FACC newly alleges that Plaintiffs' copyrights have been infringed by large language models trained by Microsoft—completely independently of OpenAI. Although Microsoft's own training of LLMs has been public since at least 2020, Plaintiffs chose not to bring claims based on those models when filing this suit. And indeed, the MDL Panel's rationale for centralizing the cases rests on the commonality of issues relating to the training of *OpenAI's* LLMs and use of them in the products offered by OpenAI and Microsoft. Now, NY Author Plaintiffs propose not only to add these unrelated models but also to completely redefine the scope of the proposed classes, changing the previous fiction and nonfiction author classes into a "Microsoft Class" and "OpenAI Class." At the same time, NY Author Plaintiffs also seek to add new *entity* named plaintiffs in the form of "loan-out" companies. Despite knowing full

well who held the rights to their copyrights when filing these lawsuits, the NY Author Plaintiffs now seek to add non-author rightsholders of a completely different character.

This case has already been litigated for 16 months under the current class definitions, which included neither models independently trained by Microsoft nor non-author entity named plaintiffs. Defendants formulated a discovery plan and strategy directed to those class definitions. The MDL Order that recently centralized competing California and NY class actions before this Court was made upon the premise that all of the cases concerned the training of OpenAI's LLMs. The competing California and New York class cases have different counsel, different class definitions, and now competing motions for leave to amend pending with radically different theories. The appointment of Interim Class Counsel pursuant to Rule 23(g)(3) has not yet occurred. Yet the NY Author Plaintiffs seek to add stale claims about unrelated Microsoft models and non-author entity rightsholders in order to upend the class definitions, disregarding the very premise upon which these actions were centralized.

In addition to adding unrelated models, a different class of named plaintiffs, and substantially altering the class definitions, the NY Author Plaintiffs also now seek to use the Proposed FACC to introduce into the case many new Microsoft products. Magistrate Judge Wang has previously denied discovery into Microsoft's recently-released generative search engine product ("New Copilot") because of the expanded discovery that it would entail. NY Author Plaintiffs now seek to go far beyond adding New Copilot to encompass a broad range of additional Microsoft products not previously in the case. Magistrate Judge Wang was correct in refusing even the more limited expansion of the case as part of the discovery process, and that refusal illustrates why this amendment should be denied.

Allowing Plaintiffs to introduce a raft of new Microsoft products into the case would necessitate not only a re-do of significant amounts of discovery, but also bring into play entirely new system architectures, new source code, new data, and many new ESI custodians. Microsoft would be required to investigate entirely new teams of researchers, scientists, engineers, and business professionals. The parties then would need time to identify custodians, negotiate search terms, review documents, produce them, and continue working through the discovery process before even getting to the point of scheduling depositions. Document discovery is nearly complete, the parties are nearly ready to commence depositions. Expanding the scope now would necessitate lengthening the schedule to substantial completion of document production by many months, again delaying the commencement of depositions, and unduly delaying a decision in this important case thus prejudicing both the Plaintiffs and the Defendants.

Moreover, there is no reason why unrelated and follow-on claims could not be brought in a separate proceeding that is guided by the outcome in this one. Indeed, to address any concerns the Plaintiffs might have had in that regard about statute of limitations issues, Microsoft offered a tolling agreement that would freeze the parties' respective positions in place as of the time the amendment was first proposed.

This is an important case. Plaintiffs seek billions of dollars in damages, and Defendants seek confirmation that they can continue to develop their groundbreaking technology. Both Plaintiffs and Defendants would benefit from a prompt decision regarding this case's central question of fair use. This is no time to be dramatically expanding the scope of this case to add multiple new theories and technologies, necessitating a clearly prejudicial tripling or more in the scope of discovery from Microsoft. The Motion should be denied.

FACTUAL AND PROCEDURAL BACKGROUND

A. Current Case Status And Schedule

1. The Instant Action

In late 2023, the Authors Guild and Alter plaintiffs filed two separate class action complaints alleging that OpenAI infringed their copyrighted works when training its large language models (LLMs) and that Microsoft was contributorily liable. *Authors Guild, et al. v. OpenAI, Inc., et al.*, No. 1:23-cv-08292-SHS ("*Authors Guild*"), ECF 1 (Sept. 19, 2023); *Alter v. OpenAI, Inc., et al.*, No. 1:23-cv-10211-SHS ("*Alter*"), ECF 1, 26 (Nov. 21, 2023 and Dec. 19, 2023). In January 2024, the Court ordered Plaintiffs to file a consolidated class action complaint. *Id.*, ECF 34. They did so, filing the First Consolidated Class Complaint ("Consolidated Class Complaint") in February 2024, alleging copyright infringement against both Microsoft and OpenAI based upon OpenAI's training of its large language models. Consolidated Class Complaint ¶¶ 412-29. Plaintiffs fall within two putative classes: nonfiction authors and fiction authors. *Id.* ¶¶ 393-94, 396-97.[1]

Discovery based on the Consolidated Class Complaint has marched along for well over a year. Tethered to the allegations in the Consolidated Class Complaint and the corresponding claims and defenses, Plaintiffs have served five sets of requests for production on Microsoft and eleven on OpenAI, as well as four sets of Interrogatories on each Defendant, and a set of requests for admission on OpenAI. Decl. of Annette Hurst ISO Opp. to Motion ("Hurst Decl.") ¶ 2. Microsoft has served four sets of requests for production on Plaintiffs and OpenAI has served nine, Microsoft has served two sets of Interrogatories on Plaintiffs and OpenAI has served four,

[1] The Court also consolidated *Basbanes, et al. v. Microsoft, et al.*, Case No. 1:24-cv-00084, Dkt. 1 (S.D.N.Y. Jan. 5, 2024) ("*Basbanes*"). The *Basbanes* Plaintiffs later dismissed their class claims to proceed individually and stayed the individual action pending a class certification decision. ECF 103 (Sept. 26, 2024).

and OpenAI has served three sets of requests for admission on Plaintiffs. *Id.* ¶ 4. In response to these discovery requests, Microsoft has produced over 82,000 documents, OpenAI has produced over 116,00 documents, and Plaintiffs have produced over 26,000 documents. *Id.* ¶¶ 3, 5. The documents produced are only a portion of the documents collected and reviewed, as Defendants finalize their review and production from a combined Court-ordered custodial hit count that exceeds 1,000,000 documents. *Id.* ¶ 6. Depositions are poised to commence. *Id.* ¶¶ 7-8.

During the discovery process, the parties have engaged in numerous meet-and-confers resulting in numerous letter requests filed with the Court. *See Authors Guild*, ECF 254, 280, 288, 317, 397. These requests, along with the ones made in the News Cases,[2] have resulted in four separate lengthy conferences before Magistrate Judge Wang. *See Authors Guild*, ECF 202, 252, 293, 321. Magistrate Judge Wang ordered two settlement conferences to facilitate the parties' agreement on a deposition protocol applicable to all of the NY cases, work that will undoubtedly be useful for the MDL. Hurst Decl. ¶ 7.

2. Related Cases and Consolidation

Prior to the two deposition planning settlement conferences, and because the California and New York class counsel were refusing to coordinate regarding depositions, on December 6, 2024, OpenAI filed a motion to centralize this case with the various other litigations pending in this Court and the Northern District of California into a single multidistrict litigation. *In re OpenAI Copyright Litigation*, Case MDL No. 3143 (N.D. Cal.), ECF 1 (Dec. 6, 2024). Each of the cases addressed in that motion commonly assert claims arising from the alleged training of only OpenAI's LLMs, with some also including claims about outputs. *Id.* That motion was

[2] "The News Cases" includes *The New York Times v. Microsoft Corporation et al.*, No. 1:23-cv-11195 (S.D.N.Y); *Daily News, LP, et al. v. Microsoft Corporation et al.*, No. 1:24-cv-03285-SHS (S.D.N.Y); and *The Center for Investigative Reporting, Inc.*, No. 24-cv-04872-SHS-OTW (S.D.N.Y).

granted on April 3, 2024, with the MDL to be centralized in this Court. Case MDL No. 3143, ECF 85; Case NYS/1:23-cv-08292, ECF 42. None of the cases involve claims arising from models trained by Microsoft rather than OpenAI.

The Court has not yet set the initial organizing conference for the MDL. Plaintiffs filed the instant Motion after the MDL transfer order was issued. But this is not the only pending motion for leave to amend in the newly centralized class cases. The California Author Class Plaintiffs also filed a Motion for Leave to Amend in the California consolidated class actions, *In re OpenAI Chat GPT Litigation*, No. 3:23-cv-03223-AMO (N.D. Cal.) ("N.D. Cal. Action"), that was previously scheduled to be argued in April but taken off calendar when the JPML issued its MDL order. N.D. Cal. Action, ECF 370. Plaintiffs in the N.D. Cal. Action seek leave to (1) add Microsoft as a defendant; (2) add a new corporate plaintiff; (3) add numerous state law claims and an *antitrust* claim; and (4) make material changes to the putative class, including expanding the relevant class period and changing the definition to enlarge the putative class. *See id.*, ECF No. 371-16 (Proposed Second Amended Consolidated Complaint). These are new claims against Microsoft, but none relate to models trained by Microsoft.

The Court has not yet appointed interim class counsel pursuant to Rule 23(g)(3) or ordered the filing of a single consolidated class complaint across the bicoastal sets of consolidated class actions.

B. Plaintiffs' Proposed FACC Would Require Re-Doing Document Review And Starting Fresh With Broad Additional Discovery.

As relevant to this Opposition, Plaintiffs' Proposed FACC seeks to expand the case in the following ways.

Unrelated Models and New Products. NY Author Plaintiffs seek to add theories concerning: (1) LLMs trained by Microsoft through Microsoft's Turing Project, which are not

alleged to have anything to do with OpenAI, and (2) newly released Microsoft products that incorporate newly added additional GPT models trained by OpenAI. *See, e.g.*, Proposed FACC ¶¶ 149-55 (Project Turing, Turing-NLG model, Megatron-Turing Natural Language Generation model); 102, 148, 178 (referring to the "successor models" of GPT-3, GPT-3.5, and GPT-4). Neither of these have ever been a part of this case, or any of the other cases in the MDL, and consequently no discovery on these technologies has yet taken place.

With respect to Microsoft's own LLMs, this appears to have been a conscious decision by the Plaintiffs as these models have been publicly known for years. In fact, Plaintiffs' cited articles about those LLMs that were published in 2020 and 2021 (see *infra* n.3), and yet Plaintiffs did not include any allegations about them when filing this action or when they filed the Consolidated Complaint.

Moreover, by adding the "successor" OpenAI models and then alleging that all Microsoft products incorporating them are at issue, *see* FACC ¶¶ 99, 102, the Proposed FACC seeks to sweep into the case many newly released Microsoft products including not only the New Copilot released late last year but many other Microsoft products such as its Office suite, Microsoft Dynamics, and the Azure platform. *See* Proposed FACC ¶ 102 ("GPT-4 and its successor models also underly Microsoft's Bing Chat product, offered through its Bing Internet search engine, and is integrated into its sales and marketing software, coding tools, productivity software, and cloud storage services.") With this deceptively short phrase, the Proposed FACC seeks to expand the case to vast swaths of Microsoft's AI business about which no discovery has been conducted so far.

New Named Plaintiffs and Class Definitions. NY Author Plaintiffs also seek leave to add additional named plaintiffs: namely, nine loan-out companies owned by some of the named

plaintiff authors. *See, e.g.*, Proposed FACC ¶¶ 48-56. Because these companies are not "authors," Plaintiffs' proposed FACC characterizes these proposed Plaintiffs as "copyright holders." *Compare* Proposed FACC ¶ 2, *with* Consolidated Complaint ¶ 2.

In addition to adding plaintiffs, Plaintiffs also seek to modify the class definitions to align with their proposed changes to add Microsoft-trained models to the case. Rather than a "fiction class" and "nonfiction class," Plaintiffs now seek to drastically re-draw the purported classes, redefining them based on the defendant that allegedly trained the model—calling the classes the "OpenAI Class" and "Microsoft Class." *Compare* Proposed FACC ¶¶ 417-18, *with* Consolidated Complaint ¶¶ 393, 396.

After receiving the Proposed FACC, Microsoft recognized that NY Author Plaintiffs might be concerned about statute of limitations issues. To allay any such concerns, Microsoft offered a tolling agreement to address the new theories in lieu of expanding the scope of these proceedings. Hurst Decl. ¶ 9 & Ex. A. Plaintiffs refused the offer without explanation.

ARGUMENT

Although leave to amend is ordinarily liberally granted, it is within the court's discretion to deny leave where there is "undue delay, bad faith or dilatory motive on the part of the movant, repeated failure to cure deficiencies by amendments previously allowed, undue prejudice to the opposing party by virtue of allowance of the amendment, [or] futility of amendment." *Foman v. Davis*, 371 U.S. 178, 182 (1962). "With respect to undue prejudice, courts consider whether amendment would: (i) require the opponent to expend significant additional resources to conduct discovery and prepare for trial; (ii) significantly delay the resolution of the dispute; or (iii) prevent the plaintiff from bringing a timely action in another jurisdiction." *Daddino v. Sanofi US Servs. Inc.*, No. 23-cv-8063, 2024 WL 1530762, at *11 (E.D.N.Y. Mar. 18, 2024), *report and*

recommendation adopted, No. 23-cv-8063, 2024 WL 1694063 (E.D.N.Y. Apr. 19, 2024) (internal quotations omitted).

The Court should deny the Motion for Leave to Amend (1) because Plaintiffs waited far too long to bring an infringement claim about Microsoft's wholly separate Turing Project and on behalf of loan-out companies they knew full well were rightsholders when filing these lawsuits, while (2) the addition of numerous additional OpenAI LLMs and the recently-released Microsoft products incorporating them would triple or more the size of this case. The Proposed FACC would substantially impair the overall administrability of the MDL by materially changing class definitions and adding matters unrelated to OpenAI as well as many new Microsoft products, burdening the Court and prejudicing Microsoft and the other parties through waste in discovery and a delay in resolution of the case. There is simply no reason these additional matters need to be litigated now as part of these centralized actions.

I. AMENDMENTS TO ADD THE "TURING MODEL" INFRINGEMENT CLAIM AND NEW PLAINTIFFS SHOULD BE REJECTED AS UNDULY DELAYED.

Plaintiffs' proposed additional infringement theory tied to Microsoft's alleged training of the Turing models as well as their addition of new entity named plaintiffs is unduly delayed and should be rejected for that reason alone. Undue delay is particularly problematic where, as here, Plaintiffs had the relevant facts at the time of suit but waited to seek amendment. *See Smith v. Westchester Cnty. Dep't of Corr.*, No. 12-cv-3941, 2014 WL 4384104, at *10 (S.D.N.Y. Sept. 3, 2014) ("[T]he Court may deny a motion to amend when the movant knew or should have known of the facts upon which the amendment is based when the original pleading was filed, particularly when the movant offers no excuse for the delay.").

Plaintiffs argue that amendment as to the Microsoft-trained models is timely because the new allegations arose from discovery produced in this case. *See* Mot. at 3-5. Yet in the Proposed

FACC, Plaintiffs cite to public blog posts from 2020 and 2021, as well as a Business Insider article from 2021,[3] as support for these new allegations. They offer no other specifics in the Motion itself. Whether or not Plaintiffs noticed this theory because of discovery material, they could have been aware of it much sooner from the very publications they cite in the Proposed FACC. Nor do Plaintiffs offer any explanation or excuse for the failure to discover these online sources by the time the complaint was filed in this action, let alone another 16 months later.

Plaintiffs also could have named their loan-out companies as plaintiffs from the outset. They knew or should have known the status of their rights and whether they had standing or needed to join others at the time they filed suit. They had access to sophisticated counsel to advise them on who should be included as plaintiffs. Instead of adding the loan-out companies, they waited to see if Defendants figured out that Plaintiffs had ownership and standing problems. There is no excuse for Plaintiffs to take a wait and see approach about a standing issue, and then seek to add additional plaintiffs of a very different character in this putative class action. The loan-out companies are entity plaintiffs whose interests may or may not be appropriately representative in this class action. They almost certainly change the nature of the claimed harm for fair use fourth-factor purposes, and they may inject significant class conflict issues 16 months

[3] *See* Proposed FACC ¶ 67 n.12 (citing *Using DeepSpeed and Megatron to Train Megatron-Turing NLG 530B, the World's Largest and Most Powerful Generative Language Model*, Microsoft Research Blog (Oct. 11, 2021)); Proposed FACC ¶ 150 n.45 (citing *Turing-NLG: A 17-billion-parameter language model by Microsoft*, Microsoft Research Blog (Feb. 13, 2020); *"Inside Microsoft's Project Turing, the team that's quietly reinventing how it develops advanced AI to move faster and take on rivals like Google,"* Business Insider, Oct. 12, 2021, https://www.businessinsider.com/microsoft-project-turing-ai-large-language-models-google-openai-2021-9); Proposed FACC 151 ¶ n.46 (citing *Turing-NLG: A 17-billion-parameter language model by Microsoft*, Microsoft Research Blog (Feb. 13, 2020), published at https://www.microsoft.com/en-us/research/blog/turing-nlg-a-17-billion-parameter-language-model-by-microsoft); Proposed FACC ¶ 153 n.48 (citing *Using DeepSpeed and Megatron to Train Megatron-Turing NLG 530B, the World's Largest and Most Powerful Generative Language Model*, Microsoft Research Blog (Oct. 11, 2021)).

into the case. Plaintiffs seek to inject potentially significant Rule 23 issues that did not previously exist, *after* Defendants have already taken a year's worth of discovery from them. Plaintiffs offer no excuse for the delay. *See Smith*, 2014 WL 4384104, at *10 (denying leave to amend where plaintiff had already been given multiple opportunities to amend and was aware of the relevant facts at the time of filing his complaint).

II. THE PROPOSED FACC WOULD PREJUDICE MICROSOFT BY EXPONTENTIALLY EXPANDING ITS DISCOVERY OBLIGATIONS AND COMPLICATING THE MDL PROCESS.

The Motion should also be denied because it would force Microsoft to conduct multiple new investigations and essentially restart document discovery just as that phase of this case is nearing its conclusion. *See Amusement Indus., Inc. v. Stern*, No. 07 Civ. 11586, 2014 WL 4460393, at *13 (S.D.N.Y. Sept. 11, 2014) ("Prejudice is particularly likely where the amendment raises new theories of recovery or would require additional discovery.") (citation omitted); *AEP Energy Servs. Gas Holding Co. v. Bank of Am., N.A.*, 626 F.3d 699, 726-27 (2d Cir. 2010) (affirming denial of leave to amend where "the impact of the proposed new" pleadings on the existing proceedings would have been "substantial" and would have required defendants to "expend significant additional resources to defend") (quotation marks omitted); *Daddino*, 2024 WL 1530762, at *11 (noting that undue prejudice can arise from "requir[ing] the opponent to expend significant additional resources to conduct discovery and prepare for trial").

Plaintiffs downplay the potential effects of the Proposed FACC, arguing that "the proposed amendment will not delay the schedule," and that Plaintiffs are "continu[ing] to push for an expeditious progression of this matter." Mot. at 10 (citing Plaintiffs' opposition to Defendants' request to continue the April 10, 2025 conference). But the practical effect of Plaintiffs' new factual allegations and legal theories is a prejudicially exponential expansion of Microsoft's discovery obligations.

A. Plaintiffs' Proposed New Theories Implicate Entirely Different System Architectures And Completely Different Groups of People From Microsoft.

Plaintiffs' new theories implicate at least two completely new system architectures: (1) Microsoft-trained research models (Proposed FACC ¶¶ 150-55); and (2) Microsoft products incorporating new OpenAI GPT models, including New Copilot (*id.*, ¶¶ 4, 99). Adding these theories against Microsoft would necessitate a restart in discovery, frustrating all prior efforts by the parties and the Court to get the case to its current position of nearly-completed document discovery.

Microsoft-Trained Models. Plaintiffs' proposed allegations seeking to introduce Microsoft-trained LLMs prejudice Microsoft at this stage of the litigation because they are entirely unrelated to the OpenAI LLMs at issue. *Amusement Indus.*, 2014 WL 4460393, at *15 ("When parties move to add claims involving completely different factual allegations and legal theories from the underlying suit, courts may exercise their discretion to deny amendment "in order to avoid unnecessary complexity and confusion.").

From their outset, these cases have been about OpenAI's training of its LLMs and the subsequent incorporation of those models into products launched by OpenAI and Microsoft. LLMs trained solely by Microsoft have never before been included in the pleadings or subject to discovery, even though their existence was no secret. Plaintiffs' efforts to include Microsoft-trained models would explode the scope of the case by requiring Microsoft to start its document discovery from scratch from new custodians and repositories that do not overlap with the discovery that has already been conducted. This problem would be exacerbated by the staleness of these allegations, which concern events that occurred prior to 2020 and 2021. Moreover, the Project Turing and the Turing-NLG and Megatron-Turing Natural Language Generation models were intended as research models and involved research collaborations with other parties not

before the Court. *See, e.g., Using DeepSpeed and Megatron to Train Megatron-Turing NLG 530B, the World's Largest and Most Powerful Generative Language Model,* Microsoft Research Blog (Oct. 11, 2021), *published at* https://www.microsoft.com/en-us/research/blog/using-deepspeed-and-megatron-to-train-megatron-turing-nlg-530b-the-worlds-largest-and-most-powerful-generative-language-model/ ("It is the result of a research collaboration between Microsoft and NVIDIA to further parallelize and optimize the training of very large AI models."). *See, e.g., Triangle Indus., Inc. v. Kennecott Copper Corp.*, 402 F. Supp. 210, 212 (S.D.N.Y. 1975) (rejecting addition of claims where they "raise[ ] the question [of] whether amendments (and supplementation) of the complaint would not add to the complications of an already over-complicated case").

Microsoft Products Based on New GPT Models. Furthermore, an innocuous-looking line in the Proposed FACC would reset Microsoft's document discovery clock to the beginning. When they bring into the case "GPT-4 and its successor models [that] underly Microsoft's Bing Chat product, offered through its Bing Internet search engine, and is integrated into its sales and marketing software, coding tools, productivity software, and cloud storage services," Plaintiffs fail to tell the Court that they seek to expand discovery to cover the entire Office Suite, Microsoft Dynamics, and the entire Azure platform. Also included in this group is New Copilot, released on October 1, 2024. Despite the similar nomenclature, New Copilot operates on an entirely different architecture and was developed by a different group of people than the deprecated Bing Chat/Copilot product, which would only "further muddle this already complex case." *Amusement Indus.*, 2014 WL 4460393, at *15. Magistrate Judge Wang already refused to add this product to the case because of the significant effect it would have on the scope of discovery. *See* ECF 293 at 3.

Moreover, to add documents relevant to the products incorporating new OpenAI GPT models, in addition to collecting new documents, Microsoft would essentially need to ***re-review*** documents that it already collected and that may have otherwise not been produced because they were not relevant to the case at the time of the previous review. Furthermore, if New Copilot is deemed within scope of this litigation, discovery would significantly increase as it would require analyzing a new set of source code, business documents, technology documents, and other material implicating an entirely new set of document custodians. The scope of such new work would exponentially increase by adding the Office suite of products, enterprise-software products, and Microsoft's cloud business. It would be like starting over on a new case that is orders of magnitude larger in discovery scope. This undertaking runs contrary to prior Court orders that have helped maintain a reasonable process for discovery. Furthermore, it simply cannot be that each time OpenAI releases a new version of GPT, it, and everything incorporating it, are to be added to the case. If so, then this case would have no end.

Plaintiffs rely on inapposite cases in support of the Motion. For example, unlike in *United States for & on Behalf of Maritime Administration v. Continental Illinois National Bank & Trust Co. of Chicago,* 889 F.2d 1248, 1255 (2d Cir. 1989), it is not the burden on Microsoft alone that warrants denial of leave to amend, but also the fact that Plaintiffs' new allegations complicate the pending MDL, as explained *infra*. And, in *State Teachers Retirement Board v. Fluor Corp.*, 654 F.2d 843, 856 (2d Cir. 1981), unlike this case, the relevant witnesses had already been deposed and the proposed amended class definition would not involve additional deponents. Here, Plaintiffs' changes would involve several groups of brand new custodians and potential deponents, both from Microsoft and from Plaintiffs, as well as discovery into additional document repositories and inspection of additional technology stacks and source code. Finally,

unlike in *Dass v. City University of New York*, No. 18-cv-11325, 2024 WL 4986914, at *5 (S.D.N.Y. Dec. 5, 2024), where the plaintiff indicated that she would not need any additional discovery, both sets of parties here would need to conduct significant additional discovery if Plaintiffs are granted leave to amend.

Microsoft has already invested significant time and resources investigating Plaintiffs' claims and conducting extensive discovery to probe their theories, and the new allegations would dramatically alter the playing field. *See Amusement Indus*, 2014 WL 4460393, at *13 (collecting cases denying motions for leave to amend when seeking to add new claims "based on different factual allegations and distinct legal theories[] from the claims already at issue in a case").

B. Microsoft Would Also Need To Propound Significant New Discovery Regarding The New Plaintiffs.

The Proposed FACC would further prejudice Microsoft because of the need for additional discovery from the loan-out companies newly named as new plaintiffs. The loan-out companies have a different set of interests in the asserted works. While both the authors and loan out companies might share copyright ownership, each group has its own nuances not relevant to the other. For example, the individual authors' ownership interests and alleged harm as the initial creators of the works would by nature differ from the "harm" the loan-out companies could claim. Furthermore, as corporate entities, the loan-out companies have nuances as class representatives that warrant being explored in discovery. The inclusion of entity plaintiffs for the first time also has the potential to significantly alter the class action framework of this case. *See Feldman v. Lifton*, 64 F.R.D. 539,543 (S.D.N.Y. 1974) (denying motion for leave to amend where amendment was "inconsistent with plaintiffs' class allegations").

C. Unnecessarily Expanding The Scope Of This Case Makes MDL Coordination Unmanageable.

Finally, there are unique concerns to this proposed expansion of the case in light of the recent centralization into an MDL, independently warranting denial on prejudice and case management grounds.

There are competing amendments from the California and New York Author Plaintiffs taking very different approaches. Across the landscape of proposed amendments, the class plaintiffs would take varying approaches by adding numerous state law claims, unrelated federal law claims (antitrust), unrelated factual theories (Microsoft Turing models), and a huge scope of recent Microsoft products. All of that is to come after this case has already been pending for 16 months. The picture presented is one of jockeying amongst counsel for position, or seeking avoidance of the MDL, rather than paying serious attention to the speedy and efficient resolution of the underlying issue of whether it is a fair use to train LLMs.

Adding Microsoft-trained models to this case would hinder judicial efficiency in the MDL. These models would come with an entirely new theory of infringement not found in any of the other cases, making management of these cases all the more complicated. *See In re Glucagon-Like Peptide-1 Receptor Agonists (GLP-1 RAs) Prods. Liab. Litig.*, No. MDL 3094, 2024 WL 5116611, at *1 (U.S. Jud. Pan. Mult. Lit. Dec. 12, 2024) (denying transfer of a case to the MDL because it would "include claims for new types of injuries" that would "significantly complicate the management of" the already complicated MDL).

Plaintiffs further prejudice Microsoft and complicate the pending MDL process by substantively changing the putative class definitions to focus away from the asserted works and onto the models those works were allegedly used to train. This change will also impede the manageability of the MDL because it would differentiate the interests of the putative classes. *See*

In re Glucagon-Like Peptide-1 Receptor Agonists, 2024 WL 5116611, at *1. In the Consolidated Complaint, the putative classes are nonfiction authors and fiction authors. Consolidated Complaint ¶¶ 392-93, 396. Plaintiffs now seek to completely shift the focus of the classes to which LLMs, either OpenAI's or Microsoft's, Plaintiffs allege the works were used to train. This change is far more than a difference in semantics—it completely alters the scope of the litigation because Microsoft-trained models have not been a part of this case, and no discovery has been conducted into them. Plaintiffs should not be allowed to change the class definitions at this stage of the litigation, particularly now that this case is part of an MDL. *See, e.g.*, *Moore v. Publicis Groupe SA & MSL*, No. 11 Civ. 1279, 2013 WL 4483531, at *13 (S.D.N.Y. Aug. 23, 2013) (denying motion to amend class definition after close of class discovery as prejudicial, where amendment would necessitate additional discovery), *report and recommendation adopted sub nom. Moore v. Publicis Grp. SA & MSL*, No. 11-cv-1279, 2013 WL 5951903 (S.D.N.Y. Oct. 30, 2013); *cf. AEP Energy*, 626 F.3d at 726-27 (affirming denial of leave to amend based on new claim and need for defendant to "expend significant additional resources").

Despite having been filed after the case was ordered centralized by the JPML, Plaintiffs' Motion makes no mention of the MDL and how the proposed amendments in class definitions would affect centralization. That is undoubtedly because the next steps should be the organization of the cases and selection of interim class counsel to represent all of the class plaintiffs, as well as the likely filing of a consolidated complaint across the class cases. *See Deangelis v. Corzine*, 286 F.R.D. 220, 223 (S.D.N.Y. 2012) ("The designation of interim class counsel is especially encouraged in cases ... where there are multiple, overlapping class actions that require extensive pretrial coordination.") (citations omitted). As a point of administrability of the MDL, only after the appointment of interim counsel should the Class Plaintiffs be

permitted to file a motion seeking leave to amend. *In re Air Cargo Shipping Servs. Antitrust Litig.*, 240 F.R.D. 56, 57 (E.D.N.Y. 2006) (summarizing the responsibilities of interim class counsel to include, *inter alia*, "making and responding to motions" on behalf of the putative class).

Here we have significant proposed changes in class definitions, and between the consolidated California and NY cases there are significant competing approaches to what claims should be included in the case. None of this should be resolved by way of this Motion while organization of the MDL and appointment of interim counsel is pending. Because the Proposed FACC would impede the efficient management of the MDL, thereby prejudicing the parties and wasting the resources of the Court, amendment is inappropriate and the Motion should be denied.

CONCLUSION

For the reasons presented herein, Microsoft respectfully requests that the Court deny Plaintiffs' Motion for Leave to Amend.

Dated: April 25, 2025

/s/ Annette L. Hurst

Annette L. Hurst (admitted *pro hac vice*)
ORRICK, HERRINGTON &
SUTCLIFFE LLP
The Orrick Building
405 Howard Street
San Francisco, CA 94105
Telephone: (415)773-5700
ahurst@orrick.com

Christopher Cariello
Marc Shapiro
ORRICK, HERRINGTON &
SUTCLIFFE LLP
51 West 52nd Street
New York: NY 10019
Telephone: (212) 506-3778
Facsimile: (212) 506-5151

ccariello@orrick.com
mrshapiro@orrick.com

Sheryl Koval Garko (admitted *pro hac vice*)
Laura Brooks Najemy (admitted *pro hac vice*)
ORRICK, HERRINGTON &
SUTCLIFFE LLP
222 Berkeley Street, Suite 2000
Boston, MA 02116
Telephone: (617) 880-1800
Facsimile: (617) 8801-1801
sgarko@orrick.com
lnajemy@orrick.com

Jeffrey S. Jacobson
FAEGRE DRINKER BIDDLE
& REATH LLP
1177 Avenue of the Americas
New York, New York 10036
Telephone: (212) 248-3191
jeffrey.jacobson@faegredrinker.com

Jared B. Briant (admitted *pro hac vice*)
Kirstin L. Stoll-DeBell (admitted *pro hac vice*)
FAEGRE DRINKER BIDDLE
& REATH LLP
1144 Fifteenth Street, Suite 3400
Denver, CO 80202
Telephone: (303) 607-3500
jared.briant@faegredrinker.com
kirstin.stolldebell@faegredrinker.com

Carrie A. Beyer (admitted *pro hac vice*)
FAEGRE DRINKER BIDDLE & REATH LLP
320 South Canal Street, Suite 3300
Chicago, IL 60606-5707
Telephone: (312) 569-1000
Facsimile: (312) 569-3000
carrie.beyer@faegredrinker.com

Elizabeth M.C. Scheibel (admitted *pro hac vice*)
FAEGRE DRINKER BIDDLE & REATH LLP
2200 Wells Fargo Center, 90 S. 7th Street
Minneapolis, MN 55402
Telephone: (612) 766-7000
Facsimile: (612) 766-1600

elizabeth.scheibel@faegredrinker.com

Attorneys for Defendant Microsoft Corporation

PROOF OF SERVICE VIA ECF

On April 25, 2025, I caused the foregoing document to be served on all counsel of record via ECF.

/s/ Annette L. Hurst
Annette L. Hurst

CERTIFICATE OF WORD COUNT

Pursuant to Local Civil Rule 7.1, I hereby certify that the foregoing document was typed using 12-point, Times New Roman font and contains 5,442 words, exclusive of the case caption, table of contents, table of authorities, signature blocks, or any required certificates, but does include material contained in footnotes or endnotes.

/s/ Annette L. Hurst
Annette L. Hurst